

知的分散OS

関 俊文, 長谷川哲夫, 原嶋秀次, 岡宅泰邦, 田村信介

(株) 東芝, システム・ソフトウェア技術研究所

知的分散システムは, 完全に集中管理機構を排したシステムである。各処理要素は, システムに与えられた目的を遂行するため, 要素間での情報交換に基ずく協調・協力メカニズムによって, 自らの役割を動的に決定する。知的分散OSは, ローカルエリア・ネットワークによって接続されたマルチマイクロコンピュータ上に開発された知的分散システムを構築するためのオブジェクト指向分散オペレーティングシステムである。本OSの特徴は, OS諸機能がオブジェクト間の通信を除いては, 個々のオブジェクトの結合によって実現されることである。本論文では, 知的分散OSの基本メカニズムを 1)オブジェクト構造, 2)オブジェクト間の情報交換機構, 3)分散ロック管理法, 4)負荷分散法, の面から述べる。さらに, 通信量の面からシステムの性能評価をする。これらの基本メカニズムを用いることにより, 容易に適用性のある大規模システムを開発, 操作することができる。

An Operating System of The Intellectual Distributed Processing System

Toshibumi SEKI, Tetsuo HASEGAWA, Syuuji HARASHIMA,
Yasukuni OKATAKU, Shinsuke TAMURA

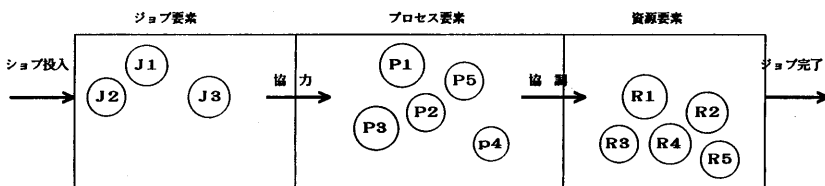
TOSHIBA corporation, Systems & Software Engineering Laboratory
70, yanagi-cho, saiwai-ku, kawasaki, kanagawa, 210 JAPAN

The Intellectul Distributed Processing System(IDPS) has no centralized managing element. IDPS system elements find their roles autonomously, in order to complete objectives given to the system. They are also able to solve conflicts among different objectives by exchanging information among elements without any supervisory element.

IDPS-OS is the object oriented distributed operating system for IDPS, which has been developed on multi-micro computers connected by local area network. The OS functions, except for communication function among objects, are realized by cooperation among objects. This paper describes the fundamental mechanisms of IDPS-OS in the following aspects: 1)object modeling, 2)message passing among objects, 3)distributed lock management, 4)load balancing. Performance evaluation results are also presented. It is expected that large scale adaptive systems can be developed and operated easier by these mechanisms.

テ エ中ブ
ス。す。ジ集オ
ヤシベ示すブ
チ、述をオすヨ
クにを造の括ジ
テ前略構類統
能キる概理種をで、
機一べて論3らこ
とア述いのれこ
造・てつム源そ
構ムいにテ資、い
のテつアス、れ、な
SSにユシス、さし
OSシスチ散セ成在
散散OK分ロ構存
分分散チのづらは
知的の分チ知、か構
知的一にブ、群機
2.1知ア1ヨト理
ム図ジク管

	Nタレ、がとにシ	がにブを作行処
	Aンさと能部た	能調才理動実、
	Lイ載く機内れ	機協は管のSて、
	の搭エいをぐ	諸・核の群〇れ
	ヤすがジベキす	の力Sサクト、
	チ示同核ブ隠続に	S協〇サクト、避
	クを、S才部手性	〇や、セエツ回
	テ造は、〇る内と張	合る。ロジ従が
	キ構にたあの内タ拡	は、結あやブー中
	の機えでト一やる	微なでやオプー集
	A算備素くドナー	的の特の理、ない
	の〇計を要エな応と	の動この處て、はへ
	S散数能ムジ要適確上	のる信ってド
	〇分複機テブ必	基造群れ通あのイ
	散的たとスオには	の構トさのでもサ
	分知れたシシ	作質現のク実間での定
	にのさいです	動的性実Sエ突もす特
	知3統エ上するム	〇ソじてクう理の
2.	図接フの作身えテ	本ブッエな管荷
2.	で一そ動自備ス	オよジ行を負



共通知識

- 機能配分
- 負荷配分

固有知識

- ロック管理
- 故障管理
- ジョブ仕様
- 機能仕様
- 処理内容

図 2 システム要素の構成

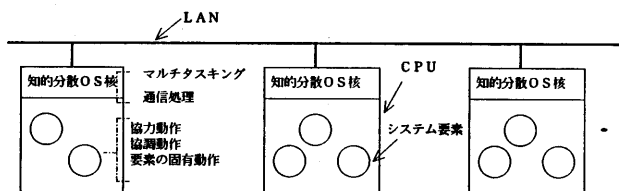


図 3 知的分散システムの物理構成

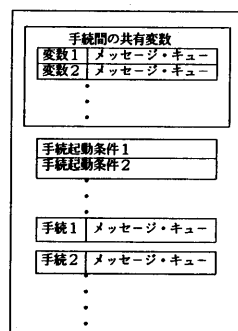


図 4 オブジェクトの構成

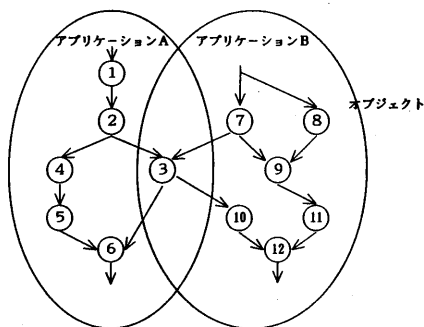


図 5 アプリケーションプログラム

るの工作。すいてはエネルギーはドップル効果で振動するが、化生機に機イナで〇さ核プのなめれ散る要し機イナの体は〇消く〇分にたさ散る主ベスムて々合にべべ解なの部うる案のkのレ〇テっ個結ら述がで々築よす提構一Sんのスよ、のさてて題け個構る現が機VOヨタシには、群みい問たが既き実ム理にシ個、合でトラうつのン加、てをズ管い、既一り、結下クグでヨ追り、施實りな、ケがおな、的ジロ質面シやよ実性ゴド少も、りての、シロ性性の一更にく諸ルイデ等べいし動い、ブンや性ケ変加なのアワマ⁴⁾アて当のなオヨ造頼りの追とム散ムははらっ担群は、シ機構信プ能、こテ分テScが図が素でび、一きやア機更るスのSc分を素要けケベ性、S変すしくしたM部化要をわとりつに、能に、〇のシ散多し、の散の理るクプ持性、性共くト中分に、現³⁾、核分定管いエがが、理とぞクを、でる実一をし特なてじる核

2.3 オブジェクトの構造

送的す統す知のすス意をいし含スレ、ント保とれ続
放動し手属通内応セ用ス正包シレハ、ライをこる
より４数に、付共ト対口に用ス修にタコ能トサしる
にに図変數３エ識、類ス共一トハル。スエト受止
核信は有変圖ジ知ブ、種ン。ピクンキンリクのを
S通ト共有。ブ通ヨ、イ。コエイでタをエジ分
〇間クの共る。オ共ジで、ラ。ジしモリドジ一部
トエ間となれ。（の）か用ス。正義とヘーブセの
は、クジきさらぞる。類るト使サ才定コンコオツ他
合エブ統続かれ種いクて成はをるイのにメの
結ジオ手手一そさのてエツ作ク係サとトラの中
のブ各とびユ。現そしジよしし現能クさ中働、のに
トオ。き及キは、実、定ブにタも層実機エ。動稼、内合
クしたる統。ジ識ては、固オと個。階てスジる。移。能ト結
エしれ手件、一知し群にスこはる、すいナブす、ら、諸エ的
ジとさ。条件有とき毎ルる識サ示用ナブ用は、が現のジ動
ブ体現に、動ツ固き統）クす知成ををスス使動な実Sブの
オ主実ウ起メと統手源た成有作係能ンラに移しく〇オ群
をによきる識手る資し生固て関機イクト問証な、たき

2.4 オブジェクト間通信

グ列をつな
ロ並る名従要か
プはあれ。必
ンいで論る。
ヨる体のれは、
シあ合手わム
一列結相なラ
ケ直の合行グ
リに群結にロ
ブよクト。的
アよくは、動
はすエ合リヨ
で示ジ結よシ
Sにブのに一
O5オトとケ
散図るクニ
分すエるプ
的は、作ジす
知ム動プ定
ラに才指て、

同協小あ結 スムなはよ 共行生通の能もで、エトゾでト都一成一
る、荷定群 シテの定に 実が共ト機て能で、ジ起ラン5ク不同生い
き合負指ト ッシ決テ機 テス効ンエ共よはのオ引ブ。ジようくれ
で場、にク よ 非ス信 ス換次ヨジ に処な、ン。ブになくつ
がるり、特エ には、デシ通 シ置副シブ。と対に方、止は、イアオと行エけ
とすよ。ジ 調に由散送 にな一オなこな更一停賃ラでがこをジ分た
こ在にる。ブ 協る自分放 ククより共うすよのう。体たオ有とす制オ使使
の存能れオ 力現で知で エエのブがよ化のS伴全ひやてA有他るでで
す数機さ名 協実間 間 ジジ次ア更るム記Oをまひいん共非た間
作が分結のる。のをトる。ト プ。の 変きラ上は止テうテオヨ 異
くト荷に数き群理クあく オオりくのでぐる合停ススこにシる。は、のクエ
なく負ト復でト管エでエ ブスよ多能処口す場のシ。シ等一い合前エ
とエツクにもクなジ欠ジ ヲセに 機対ブ対の体はい。御理ケて場名ジ共生
こジ持エ時とエドブ可。る。ジロち即するでムにこ全更な制処りしるつブ
すのブのジ同こエドオ不オす。ブこ即すみテ求、ム変は理ンブ有じ持オ共生
すオ構。る。ブワ全がを現たる。用のヌ要が、テのと管ヨア共生を
識の機オば、すオムの合れ実ますする。使更シ更るストこのシ。をが能て、
意名調のれ合 テ内結こり 有さじに変を変あシクす等クは、3合機し

デクに起と接能例數側や般ジジそ
とエれて之間機信多きキ一プー
形ジそつる・S受、送統、オセい、
出ブがよい接○とく、手いはッ多
呼オトに用直な側なで、きなグメが
きクとををう信はのべらんらと
続え、エこ懸問よ送でるす限ミかこ
構手備ジる動クトのジ合な出はイ群
機をブす形クトの一結と呼とタトれ
驅は、態オ信驅エ協一合呼とタトれ
驅ト形る送なジ、ッ純結のききエ定
トク動あをうブ力、メ単のトで統ジ決
クエ驅、ジよオ協はの群ク定手ブでで
エジ兩は、一のもの、でトエ指きオ障
ジブのききこ數る、し対クジをべ連段
ジオ形統ッ、複き際トエブジそ関た
オの出手メる、でにたジオン動がは
ッ々呼のてれっ合行ッ手ミ驅ト信
２個タ内しきよ結実いオ相いはク受
一ト對動にのとのがタにエを

Figure 1 is a diagram illustrating the execution of a program with two objects, OBJ_A and OBJ_B, and a third object OBJ_C. The diagram shows the state of the program where OBJ_A is waiting for data from OBJ_B.

OBJ_A contains the following steps:

- MA
- send(B, MB, ..., mes)
- wait(X)
- Y=X+Y

OBJ_B contains the following steps:

- MB
- X=1
- mes=X
- send(A, X, ..., mes)

OBJ_C contains the following step:

- MC

Arrows indicate the flow of control and data:

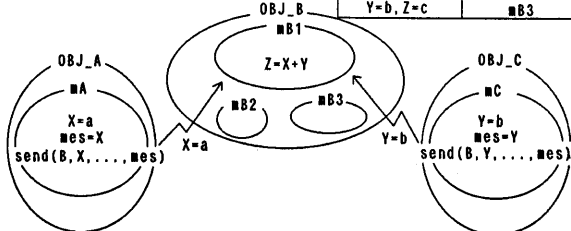
- An arrow labeled "手続呼出し" (Procedure call) points from MA in OBJ_A to MB in OBJ_B.
- An arrow labeled "データ呼出し" (Data call) points from wait(X) in OBJ_A to X=1 in OBJ_B.

The diagram illustrates the state where OBJ_A is waiting for data from OBJ_B.

受信アドレス	実行時刻	無効時刻	本 文
--------	------	------	-----

サイト名	要素名	要素内アドレス (手続名/変数名)
------	-----	----------------------

起動条件	起動メソッド
$X=a, Y=b$	■B1
$X=c, Z=a$	■B2
$Y=b, Z=c$	■B3



<4>

な係いすトブ位わらな”に”にい工”せしるよ送オな
行間高定クオ上表フ”盤作方、下てい、盤ツ信ある。じなく
に層の指エは最を”ン施工一のれな合、旋ム受でる。ジなが
的階層をジ〇”と一”に”械続かい、同”クさせ要と
率に階前ブ1る。械”レ”に”械続かい、同”クさせ要と
効名は名オ図あ機盤ぐる。下て機接わたと、に作エ送ツ必こ
係をトト）い”で”旋”いのれ”でがしるトエジ転ゑる
関信クくるなる。例な”とて”さは名前信すク”ブにに取れ
層通エエあらき一”的る”し械続”ト名送信ジ”オ”域けさ
階送ジジの知でての般あ車に機接ンクをを送ジち、位盤傾受断
名放ブジ性もが造一に送存”で一エトジでブ即下スたを中
トはオオ般前と構は位盤がは名レジク一宛オ”のいれジが
クでに、一名こ層名下”群”トクブエセ名”る。そらさ一理
エSらるで、る階下の、ト盤ク”オジッ象盤れはフ定セ処
ジ〇さいの、すのクそびクスエとるブメ抽スさジ”限ツの
ブ散、て象より合名エ、及エイジ”なオにるイ送一とがメト
オ分の、し抽に結トエ、”ジラブ車”る群なウ放セ”困、ク
）的め入りともクブお盤ブフオ送動あ械”フがツ盤範れ、エ
3知た導よこエオてスオ”る搬移、機作”ジメ旋てさジ
うを（る群ジのしイるとな”る。作工と一た”っ信づる。

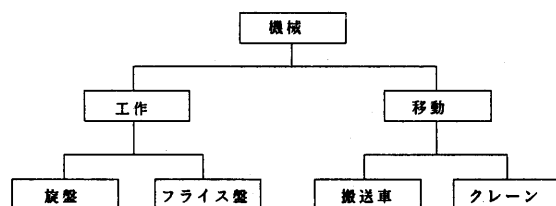


図 10 システム要素名の階層

(ジ)分すツラ行求実要なるをドキリこし信
ブ智知協だ信た行し、特に要ウス刻するは消
オ(サ)ん通つ突過り、う行ラチ時棄取刻削
理ッる力、含はなめ超よが、よ実ル在破の時
管セす協をとにたをに理るトベ、マ現を求効働
の口現が求こそ要の刻と処えくらり、ト要無
サブ実群要の不こ時こし行エなよ際、ン行
ッるをト行ごで、効る消くジに割のべたれ
セよ信ク実、等る無すりなブ列割そイたる
ロに通エのが、障す視取と才行分、つわわ
ブ割間ジ定る故にし、無のこはち時る持なな
は分トブ不れア難加が求める核待くいをに行
能時クオ先さエ困付核要けSてずて刻要に
機のエ、て換有ををS行設Oし基し時不的
のヘジは、あ交D理刻O実をはとに現効、動
核群ブで、があ一処時をた構にと法実無し、自
ストオS程ジハの効求っ機動的ンをいよがに
OK、O過一や求無要な理体ベビグさに理
エと、散るせ一要に行に管具イロン小と処

3. 自調機せず、協算関オさ

[illegible]

能シ速決用シム野
機的高解速驗イ分
なれの等ら実たる
的さ信上かけアれ
律成通向掛けアさ
た。自構能範向リ求
つからは、性なスに要
なにかにの能イ更が
うに合めム可フ答
うよ集たズ、オガ、応
よのるるリが、る速る。
るムなれごい、在、い高い
きて、単さル多現てのて
で、スの化アは、し等え
がシ素用散点る。策ム考
と散要実分題いて構テ考
をこたはて種問いてス用
るのつ、各ききめムシ適
め知持テ、ベ進テ御の、
高をス化、すをス制へ

- [1] 計測自動制御学会 "分散と協調特集", 26(1) (1987)
- [2] 田村, 他: "知的分散システムのアーキテクチャ", 電学誌 108-C(6) (1988)
- [3] Andrew S. Tanenbaum et al. : "Distributed Operating Systems", Computing Surveys, 17(4) (1985)
- [4] M. Accetta et al. : "MACH: A New Kernel Foundation For UNIX Development", Proc. of USENIX'86, 93/112 (1986)
- [5] J. H. Howard Jr : "Mixed Solutions for the Deadlock Problem", Communications of the ACM. 16(7) 427/430 (1973)
- [6] R. C. Holt : "Some Deadlock Properties of Computer Systems", Computing Surveys, 4(3) 179/196 (1972)
- [7] K. M. Chandy et al. : "Distributed Deadlock Detection", ACM Trans. Computer System, 1(2) 144/156 (1983)
- [8] S. S. Isloor et al. : "An Effective ON-LINE Deadlock Detection Technique for Distributed DBMS", Proc. of COMPSAC'78 283/288 (1978)
- [9] S. Tamura et al. : "Distributed Deadlock Avoidance and Detection in Intellectual Distributed Processing System", Proc. of SMC'88, 1279/1282 (1988)
- [10] W. Whitt: "The Queueing Network Analyzer", The Bell System Technical Journal, Nov. 2779/2843 (1983)
- [11] Yoneda et al. : "Evaluation of a queueing network package", Proc. of SCSC'88, 694/699 (1988)

Figure 1 is a line graph showing the relationship between the number of workstations (WS) and response time (レスポンスタイム) and utilization (利用率).

The X-axis represents the number of workstations (WS数), ranging from 0 to 40. The left Y-axis represents response time (レスポンスタイム) in seconds (sec), ranging from 0 to 140. The right Y-axis represents utilization (利用率), ranging from 0.0 to 1.2.

Three lines are plotted:

- WSの処理時間 (含むLANの処理時間)** (WS processing time including LAN processing time): This line shows a steep, nearly linear increase in response time as the number of workstations increases.
- WSの処理時間** (WS processing time): This line is nearly identical to the first line, showing a steep, nearly linear increase in response time.
- LANの利用率** (LAN utilization): This line shows a much slower, nearly linear increase in utilization as the number of workstations increases.

A vertical dashed line is drawn at WS数 = 35, indicating the point where the response time reaches approximately 115 seconds and the LAN utilization is approximately 0.05.

図 14 WS 数とレスポンスタイム（利用率）の関係